DOCUMENT RESUME

ED 089 384 CS 500 643

AUTHOR Hensley, Wayne E.; Batty, Paul W.
TITLE A Comparison of Instructor and Peer Evaluations of Student Speeches in a Public Speaking Course.
PUB DATE Feb 74
NOTE 13p.; Paper presented at the Conference of the Directors of the Basic Speech Communication Course of Midwestern Universities (Boulder, February 7-9, 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS College Students; *Educational Research; Evaluation Criteria; *Evaluation Methods; Evaluation Techniques; *Grading; *Public Speaking; Speech Evaluation; *Student Evaluation

ABSTRACT

The purposes of this study were to examine the relationship of students' and instructors' evaluations, to discover which elements of evaluation seem to best predict the grade awarded by each group, and to compare the dimensions of evaluative judgments for students and instructors. College students enrolled in a public speaking course at Indiana University served as subjects. The public speaking course was designed so that subjects presented speeches to classes with which they were unfamiliar. On days when students were not scheduled to speak, they remained in their regular class to serve as an audience. Each speaker received a written critique and grade from all the student audience members as well as from a visiting instructor. A five point rating scale was used by all raters to record the judgments of the speeches. Some of the results indicated that student audiences pay more attention to general effect and adaptation than do instructors in assigning grades. Also, both students and instructors come to highly similar conclusions concerning student speeches, but evidently for different reasons. (WR)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

COMMENTS WELCOME

ED 089384

A Comparison of Instructor and Peer Evaluations of Student Speeches in a Public Speaking Course

Wayne E. Hensley and Paul W. Batty

Indiana University

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Wayne E. Hensley
Paul W. Batty
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

We are indebted to Candace Brown and Diane Tobin who assisted in data preparation; Sonya Blixt and Candace Carrett offered valuable statistical advice.

Paper presented at the 1974 Conference of the Directors of the Basic Speech Communication Course of Midwestern Universities; Boulder, Colorado, February 7-9, 1974.

CS 500 643

Both instruction and learning in public speaking courses depend heavily on the ability to evaluate speech making. Teachers need to identify effective speeches; students must have some means of assessing their progress. Whatever the definition of success, one fact is clear: no speaker succeeds unless an audience thinks so. Speech success depends on effectiveness with an audience; no public speaker, to our knowledge, ever swayed masses in an empty room. Even consummate practitioners of oral communication are sometimes denied success--witness Clarence Darrow in Hawaii. Since success is an audience- centered variable, it is worthwhile to ask on what basis audiences evaluate speeches. This and related questions form the central focus of the present paper.

The disturbing element about classroom speech evaluation is that we seem to lose touch of the concept of "audience" when designing rating blanks. Although we teach audience analysis and adaptation, our judgments of speech effectiveness have most often been made in terms of an audience of <u>one</u>: the instructor. Even when evaluations have been gathered from more heterogenous audiences, their worth has been discussed in terms of agreement with an instructor's rating.

However, an important question remains unanswered. There are more students present in the classroom than there are instructors. True audience adaptation by the speaker must take these listeners into account. If we believe that public speaking is communication, it is appropriate to ask if the ratings of instructors accurately reflect the evaluation of the audience to whom the speech may be

delivered. In other words, are there systematic and definable relationships between the judgments made of a public speech by each of the two audiences--the instructor and the students?

Previous Research

Several previous investigations have examined the relationship between instructor and student evaluations of public speeches. One of the earliest of these was done by Monroe, et. al. (1935) comparing instructor-awarded grades and student ratings of the same speeches. The total picture of student and instructor evaluations led these researchers to conclude:

> "instructors' grades and students' ratings agree fairly closely as to the relative effectiveness of different speakers, and to that extent students' ratings may be said to be valid measures of judged effectiveness of the same type recognized as such by speech instructors." (pp. 25-26)

In short, the conclusion was that the students and the instructors, using the same basis for judgments, arrive at highly similar conclusions concerning speakers.

In an unpublished Master's thesis, Bowers (1959) compared the grade awarded to student speakers by groups of teachers ($N \approx 3$) and students ($N \approx 10$). Examining the relationship at the end of two separate semesters of a public speaking course, the observed correlations were found to be .56 and .61. Wiseman and Barker (1965) found no differences between composite instructor ratings and the average of five composite student ratings. However, they did find that--overall--students tend to award higher grades than instructors.

Finally, Shepard (1970) discovered that groups of students evaluate speakers into the same rank-order as do instructors.

In sum, the research evidence seems to point out that while students may grade a bit higher than do instructors, the relative order of those grades is highly consistent with the order perceived by instructors. However, in all these studies, we still do not know whether the basis upon which the students grade is the same as it is for instructor's grades. Monroe's assertion (1935) remains an assertion. Further, we do not know to what extent the components or elements of evaluation predict the final grades awarded.

In light of these issues, the purposes of this study are threefold: (1) to examine the relationship of students' and instructors' evaluations, (2) to discover which elements of evaluation seem to best predict the grade awarded by each group, and (3) to compare the dimensions of evaluative judgments for students and instructors.

Procedure

The public speaking course at Indiana University is a two-hour semester course designed to improve each students' ability to address an audience. Often the student becomes so thoroughly familiar with the other members of the class that audience adaptation, at least near the end of the course, becomes almost a reflex. Consequently, at least one opportunity is provided for each student to speak to an "unfamiliar" audience. During the last two weeks of the course, the final speech is delivered to, and evaluated by, a class other

than the one normally attended by the student. From the perspective of the student, the regular class is exchanged for another class and instructor.

The "exchange" speech is accomplished by having student speakers from one class--usually 5 at a time--go to another section of the same course meeting at the same day and hour. Simultaneously, 5 of the 20 students in the other class are coming to visit the first class. Thus, from the student speaker's viewpoint, the audience is comprised of the 15 or so students in the visited class plus the new instructor. On the days when students are not scheduled to speak, they remain in their regular class to serve as an audience. Each student speaker receives a written critique and grade from all the student audience members as well as from the visited instructor. Usually, the grade for this speech is some combination of student and instructor awarded grades. It should be emphasized that the exchange speech is a regular and quite well known feature of the course. When students in this course make speech evaluations, they are performing an expected and meaningful element of the course. Thus, any element of the "laboratory" atmosphere which occasionally contaminates investigations, may be discounted in this study.

The course itself, in the Fall semester of 1973, was taught by one Ph. D. (in this case the senior author) and by 17 graduate student Associate Instructors, all but 3 of whom held the Master's degree in Speech. Each of the 48 sections of the course had an enrollment of from 18 to 20 students.

In scheduling the exchange speeches all students were randomly assigned both to the day of presentation and to speaking order on the day when they spoke. The sample of speech evaluations for this study was drawn only from the second day of the exchange speeche Sampling this one day reduced the possible number of evaluations but it assured that student evaluators had at least one day of experience with the evaluation forms and with the general format of the assignment itself.

All instructors and student evaluators used a common rating form to record their judgments of the speeches. The form was composed of a 5-point Likert type scale of the following items: "Thesis," "Adjustment to Audience," "Supporting Material," "Organization," "Presentation," and "General Effect." In addition, a letter grade was awarded to the speaker by all students and the instructor. These 7 evaluations formed the basis for the study.

Results and Discussion

The first concern was to examine the relationship between the instructor awarded grades and the student awarded grades. Usable data for 70 such comparisons was available. Since, in some instances, the grades awarded by students were badly skewed, the median class grade was compared with the instructor's grade. The correlation between these grades was .56 ($p < .01$). (This correlation is remarkably close to those found by Bowers (1959) --.56 and .61.) Our finding strongly supports the fact that instructors' grades and students' grades covary together.

The next concern was to determine which of the six scales of evaluation would best predict awarded grades by students and by instructors. To accomplish this, two stepwise regressions were performed. It was decided not to retain any variable in the regression unless it added at least 1% explained variance to the solution. Using this criteria, the results are shown in Table 1.

Insert Table 1 about here.

Keeping in mind that the two problems were solved in precisely the same manner and the variables entered in the same order, the two solutions differ substantially. General Effect is the most explanatory variable in the student solution; it does not even appear in the instructor solution. Organization adds 13% explained variance for instructors; it is totally absent for students. Whatever is happening in the evaluative process, it appears to be very different for instructors and for students.

This result prompted us to examine the respective factor structure for students and instructors. (A standard orthogonal rotation was chosen; Multiple R^2 was inserted in the diagonals.) The results are shown in Table 2.

Insert Table 2 about here.

In both cases, the result was a clear one factor solution.

This finding partially corroborates Becker's (1962) findings. He found 3 factors but they all appeared to be measuring one primary concept. In the present study, each group seems to use the scales as measures of "Speaking Effectiveness." However, it is worth noting that the factor loadings differ for students and instructors indicating the items differ in importance for each group. If we examine the rank order correlations between factor loadings for both groups, Spearman's rho reveals there is no systematic relationship between the two solutions ($r_s = .54$; p is not significant). We can only conclude that there is no relationship between the relative importance of the same scale items for students and instructors.

Finally, our concern was to compare the dimensions of evaluation by students and instructors. Desiring to find those variables which maximally distinguished between the two groups, factorial discriminant analysis was the tool chosen (Cooley and Lohnes, 1971). All six scales plus awarded grades were entered into the problem. Table 3 shows the results.

Insert Table 3 about here.

All of the seven variables figure in the discriminant function. It is interesting that one of the <u>least</u> important of these is the awarded grade. The most important is the way in which the scale of organization is used by the two groups. The other scales fall

someplace in between. In short, both students and instructors come to highly similar conclusions concerning student speeches--but for different reasons.

Implications

Our discussion of the students in the class as the audience should not be taken to mean that instructors' ratings are of no value. The observations reported here suggest that student audiences pay more attention to general effect and adaptation than do instructors in assigning grades. Their ratings may be a more accurate indication of the speaker's "success" with the audience. Nonetheless, the instructor's evaluation, which focuses on presentation and organization, may be of more value in helping the student to improve in public speaking. Each type of evaluation should be used for its own best function.

On another front, we would urge caution in adopting peer evaluation as a substitute for a teacher in these days of vanishing budgets. Those who might be tempted to advocate peer evaluation because "you get the same answers anyway" are probably correct--but for very different reasons than first suspected. Surely, one study cannot answer all the questions concerning evaluation. However, one finding seems rather clear. The evaluation of student speakers by other students and by instructors may be very similar; but, the answers are reached for quite different reasons.

References

Becker, Samuel L. "The Rating of Speeches: Scale Independence." Speech Monographs 29 (March, 1962) 38-44.

Bowers, John Waite. "A Comparison Between Speech Evaluations by Groups of Speech Teachers and by Groups of Student Judges." Unpublished Master's Thesis, University of Kansas, 1959.

Cooley, William W. and Paul R. Lohnes. Multivariate Data Analysis, New York: John Wiley & Sons, 1971.

Monroe, A. H., H. H. Remmers and Elizabeth Venemann-Lyle, "Measuring the Effectiveness of Public Speech in a Beginning Course." Bulletin of Purdue University, Studies in Higher Education, No. 29, 1936.

Shepherd, David W. "Students Judging Student Speeches." Central States Speech Journal 21 (Fall, 1970) 196-198.

Wiseman, Gordon and Larry Barker. "A Study of Peer Group Evaluation." Southern Speech Journal 31 (Winter, 1965) 132-138.

Table 1

Regression of Six Evaluative Scales on Grades Awarded

Variable	R	Increase in Explained Variance	Total Explained Variance
Students (N = 456)			
General Effect	.732	----	.536
Presentation	.796	.097	.633
Supporting Material	.818	.035	.668
Adjustment to Audience	.835	.028	.697
Thesis	.841	.011	.708
Instructors (N = 74)			
Presentation	.743	----	.552
Organization	.825	.129	.680
Supporting Material	.847	.037	.717
Thesis	.861	.023	.741

Table 2

Unrotated Factor Solutions

for Six Evaluative Scales

(Students and Instructors)

Variable	Students (N = 502)	Instructors (N = 74)
Thesis	-.628	.573
Adjustment to Audience	-.704	.641
Supporting Material	-.608	.631
Organization	-.684	.676
Presentation	-.729	.702
General Effect	-.765	.639

Table 3

Discriminant Analysis of Six Evaluative Scales and Grades for Instructors and Students

Variable	Discrimination Index*
Grades	.566
Thesis	.395
Adjustment to Audience	.744
Supporting Material	.592
Organization	.815
Presentation	.738
General Effect	.688

*These coefficients may be thought of as betas or as factor loadings in their function. They are measures of how important the variable is in distinguishing between the two groups. see Cooley and Lohnes (1971) for further discussion.